

What is claimed is:

1. A software program for providing instructions to a processor which controls a system for applying actor-critic based fuzzy reinforcement learning, comprising:
  - a database of fuzzy-logic rules for mapping input data to output commands for modifying a system state; and
  - a reinforcement learning algorithm for updating the fuzzy-logic rules database based on effects on the system state of the output commands mapped from the input data, and
  - wherein the reinforcement learning algorithm is configured to converge at least one parameter of the system state towards at least approximately an optimum value following multiple mapping and updating iterations.
2. The software program of Claim 1, wherein the reinforcement learning algorithm is based on an update equation including a derivative with respect to said at least one parameter of a logarithm of a probability function for taking a selected action when a selected state is encountered.
3. The software program of Claim 2, wherein the reinforcement learning algorithm is configured to update the at least one parameter based on said update equation.

4. The software program of any of Claims 1-3, wherein the system includes a wireless transmitter. 3

5. A method of controlling a system including a processor for applying actor- critic based fuzzy reinforcement learning, comprising the operations:

- mapping input data to output commands for modifying a system state according to fuzzy-logic rules;
- updating the fuzzy-logic rules based on effects on the system state of the output commands mapped from the input data; and
- converging at least one parameter of the system state towards at least approximately an optimum value following multiple mapping and updating iterations.

6. The method of Claim 5, wherein the updating operation includes taking a derivative with respect to said at least one parameter of a logarithm of a probability function for taking a selected action when a selected state is encountered.

7. The method of Claim 6, wherein the updating operation includes updating the at least one parameter based on said derivative.

8. The method of any of Claims 5-7, wherein the system includes a wireless transmitter. 3

- ✓ 9. A system controlled by an actor-critic based fuzzy reinforcement learning algorithm which provides instructions to a processor of the system for applying actor-critic based fuzzy reinforcement learning, comprising:
- the processor;
  - at least one system component whose actions are controlled by said processor;
  - at least one storage medium accessible by said processor, including data stored therein corresponding to:
    - a database of fuzzy-logic rules for mapping input data to output commands for modifying a system state; and
    - a reinforcement learning algorithm for updating the fuzzy-logic rules database based on effects on the system state of the output commands mapped from the input data, and
    - wherein the reinforcement learning algorithm is configured to converge at least one parameter of the system state towards at least approximately an optimum value following multiple mapping and updating iterations.
10. The system of Claim 9, wherein the reinforcement learning algorithm is based on an update equation including a derivative with respect to said at least one parameter of a logarithm of a probability function for

taking a selected action when a selected state is encountered.

11. The system of Claim 10, wherein the reinforcement learning algorithm is configured to update the at least one parameter based on said update equation.
12. The system of any of Claims 9-11, wherein said at least one system component comprises a wireless transmitter.

11/03/2017